

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-189570

(P2002-189570A)

(43) 公開日 平成14年7月5日(2002.7.5)

(51) Int.Cl. ⁷	識別記号	F I	テ-マ-ト*(参考)	
G 0 6 F 3/06	3 0 4	G 0 6 F 3/06	3 0 4 E	5 B 0 1 8
	3 0 1		3 0 1 X	5 B 0 6 5
	3 0 5		3 0 5 G	5 B 0 8 2
12/00	5 3 3	12/00	5 3 3 A	
12/16	3 1 0	12/16	3 1 0 M	
審査請求 未請求 請求項の数 8 O L (全 14 頁)				

(21) 出願番号 特願2000-391833(P2000-391833)

(22) 出願日 平成12年12月20日(2000.12.20)

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 中村 泰明

神奈川県小田原市国府津2880番地 株式会社日立製作所ストレージシステム事業部内

(72) 発明者 田淵 英夫

神奈川県小田原市国府津2880番地 株式会社日立製作所ストレージシステム事業部内

(74) 代理人 100075096

弁理士 作田 康夫

最終頁に続く

(54) 【発明の名称】 記憶システムの二重化方法および記憶システム

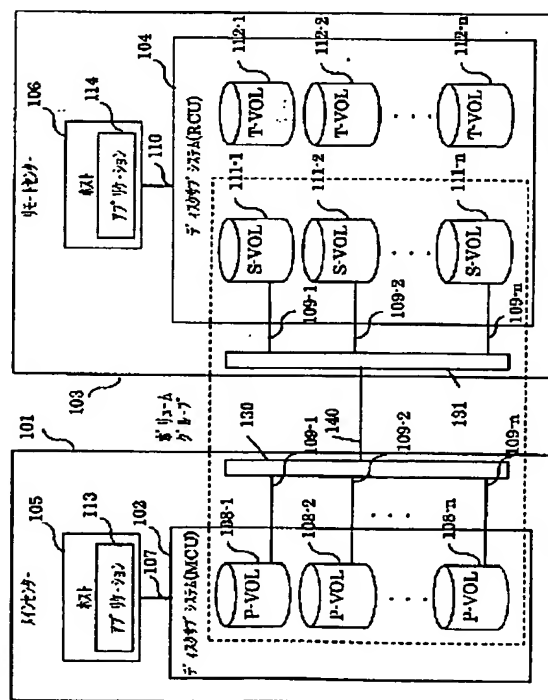
(57) 【要約】

【課題】 リモートコピー機能において、データ内容の一貫性を保持するグループとして定義されたペア論理ボリューム群において、前記グループ内の全ペア論理ボリュームの二重化を完了するまでの時間を削減し早期二重化を図る。

【解決手段】 コピーを一時停止するような事象が発生した場合、一時停止の原因が解消してから前記グループ内のペア論理ボリューム群を一度に二重化を開始するかわりに、一時停止の原因と関係ない前記グループ内のペア論理ボリュームから順次二重化を実施するようにする。

【効果】 一時停止の原因に関係しない、前記グループ内の全ペア論理ボリュームの二重化を完了するまでの時間を削減できる。

図 1



【特許請求の範囲】

【請求項1】第1記憶システムが有する複数の論理ボリュームのデータを、第2記憶システムへコピーする記憶システムのデータ二重化方法であって、
前記複数の論理ボリュームのデータを前記第2記憶システムへコピーする第1ステップと、
前記複数の論理ボリュームのデータの前記第2記憶システムへのコピーを中断する第2ステップと、
前記複数の論理ボリュームのうち、1つ以上の論理ボリュームのデータを、前記第2記憶システムへコピーを開始する第3ステップと、
前記複数の論理ボリュームのうち、前記1つ以上の論理ボリューム以外の論理ボリュームのデータを、前記第3ステップに遅れて前記第2記憶システムへコピーを開始する第4ステップとを有する記憶システムのデータ二重化方法。

【請求項2】前記コピーされたデータが、前記複数の論理ボリュームに対応する前記第2記憶システム内の複数の論理ボリュームに書込まれる第5ステップを有し、
ここで、前記第1記憶システム内の前記各論理ボリュームへの前記データの書込み順序と、前記第2記憶システム内の前記各論理ボリュームへのデータの書込み順序が一致している請求項1記載の記憶システムのデータ二重化方法。

【請求項3】第1記憶システムが有する複数の論理ボリュームのデータを、第2記憶システムへコピーする記憶システムのデータ二重化方法であって、
前記複数の論理ボリュームのうち、ある論理ボリュームのデータを、前記第2記憶システムへコピーを開始する第1ステップと、
前記複数の論理ボリュームのうち、前記ある論理ボリューム以外の他の論理ボリュームのデータを、前記第1ステップに遅れて前記第2記憶システムへコピーを開始する第2ステップとを有する記憶システムのデータ二重化方法。

【請求項4】前記コピーされたデータが、前記複数の論理ボリュームに対応する前記第2記憶システム内の複数の論理ボリュームに書込まれる第3ステップを有し、
ここで、前記第1記憶システム内の前記各論理ボリュームへの前記データの書込み順序と、前記第2記憶システム内の前記各論理ボリュームへのデータの書込み順序が一致している請求項3記載の記憶システムのデータ二重化方法。

【請求項5】前記第1ステップは、前記ある論理ボリュームのコピーを開始する前に、前記ある論理ボリュームのデータがコピーされるときに使用されるデータ転送機器の状態を参照する請求項3記載の記憶システムのデータ二重化方法。

【請求項6】前記第2ステップの後、前記第2記憶システムへコピーされたデータの複製を前記第2記憶システ

ム内に作成するステップを有する請求項3記載の記憶システムのデータ二重化方法。

【請求項7】第1記憶システムが有する複数の論理ボリュームのデータを、第2記憶システムへコピーする記憶システムのデータ二重化方法であって、
前記複数の論理ボリュームのデータを前記第2記憶システムへコピーする第1ステップと、
前記複数の論理ボリュームのデータの前記第2記憶システムへのコピーを一時中断する第2ステップと、
前記第2ステップの後、前記第2記憶システムへコピーされたデータの複製を前記第2記憶システム内に作成する第3ステップと、
前記複数の論理ボリュームのうち、1つ以上の論理ボリュームのデータを、前記第2記憶システムへコピーを開始する第4ステップと、及び、
前記複数の論理ボリュームのうち、前記1つ以上の論理ボリューム以外の論理ボリュームのデータを、前記第4ステップに遅れて前記第2記憶システムへコピーを開始する第5ステップとを有する記憶システムのデータ二重化方法。

【請求項8】第1記憶システムと、前記第1記憶システムに接続され前記第1記憶システムに格納される複数の論理ボリュームに含まれるデータの複製を記憶する第2記憶システムを有する記憶システムであって、
複数の論理ボリュームが格納される前記第1記憶システム内のディスクドライブと、
前記複数の論理ボリュームのデータが転送される複数のデータ転送機器と、
前記複数のデータ転送機器の状態を記録するメモリと、
及び、
前記メモリに記憶された状態に基づいて前記複数の論理ボリュームのデータを前記第2記憶システムへコピーする状態を管理する前記第1記憶システム内のプロセッサとを有し、
ここで、前記プロセッサは、複数の論理ボリュームのデータの複製を行う際、前記状態に基づいて、前記複数の論理ボリュームのうちの一部の論理ボリュームを他の論理ボリュームに先行してコピーを開始することを特徴とする記憶システム。

【請求項9】前記第2記憶システムは、前記第1記憶システムに接続され前記第1記憶システムに格納される複数の論理ボリュームに含まれるデータの複製を記憶する第2記憶システムを有する記憶システムであって、
複数の論理ボリュームが格納される前記第1記憶システム内のディスクドライブと、
前記複数の論理ボリュームのデータが転送される複数のデータ転送機器と、
前記複数のデータ転送機器の状態を記録するメモリと、
及び、
前記メモリに記憶された状態に基づいて前記複数の論理ボリュームのデータを前記第2記憶システムへコピーする状態を管理する前記第1記憶システム内のプロセッサとを有し、
ここで、前記プロセッサは、複数の論理ボリュームのデータの複製を行う際、前記状態に基づいて、前記複数の論理ボリュームのうちの一部の論理ボリュームを他の論理ボリュームに先行してコピーを開始することを特徴とする記憶システム。

【請求項10】前記第2記憶システムは、前記第1記憶システムに接続され前記第1記憶システムに格納される複数の論理ボリュームに含まれるデータの複製を記憶する第2記憶システムを有する記憶システムであって、
複数の論理ボリュームが格納される前記第1記憶システム内のディスクドライブと、
前記複数の論理ボリュームのデータが転送される複数のデータ転送機器と、
前記複数のデータ転送機器の状態を記録するメモリと、
及び、
前記メモリに記憶された状態に基づいて前記複数の論理ボリュームのデータを前記第2記憶システムへコピーする状態を管理する前記第1記憶システム内のプロセッサとを有し、
ここで、前記プロセッサは、複数の論理ボリュームのデータの複製を行う際、前記状態に基づいて、前記複数の論理ボリュームのうちの一部の論理ボリュームを他の論理ボリュームに先行してコピーを開始することを特徴とする記憶システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、コンピュータが参照及び更新するためのデータを格納する記憶システムに係わり、特に記憶システムの保有するデータを二重化する方法及びその記憶システムに関する。

【0002】

【従来の技術】メインのデータセンターで保有するディスクサブシステム（記憶システム）のデータは、地震等の災害により消失する恐れがある。この消失に備え、遠隔地のリモートセンターのディスクサブシステム（記憶

システム)に、そのデータのコピーを作成し、即ち、データを二重化しておくことで、このデータ消失をリカバーする方法がある。そして、このデータを二重化する方法は、いわゆるリモートコピー機能により既にいくつか実用化されている。

【0003】リモートコピー機能は、メインセンターが保有するデータの内容をリモートセンター側に、単にある時点のデータ内容としてバックアップする機能を提供するだけでなく、メインセンター内のホストコンピュータ(上位装置)からディスクサブシステムにデータの更新(書き込み)指示があったとき、リモートセンターのディスクサブシステムに対しても、当該更新データを書き込む機能である。これにより、メインセンターのシステムで何らかの障害が発生し、メインセンターにあるディスクサブシステムのデータが使用不可能になった場合、リモートセンターのディスクサブシステムが保有するデータを使用することで、即座にメインセンターでの処理を引き継ぐことができる。

【0004】この場合にリモートセンター側のディスクサブシステムでは、メインセンターでの処理を引き継ぐために、障害時のメインセンター側のデータと整合性を持ったデータを保有している必要がある。言い換えればメインセンターでのデータ更新の順序と、リモートセンターの更新順序が整合していなければならない。

【0005】このような更新順序の整合性を保持する技術がいくつか公開されている。例えば特開平6-290125および特開平11-85408によるメインセンターのディスクサブシステムの更新処理とは非同期にコピーを行うリモートコピーのシステムでは、データに付与された時刻情報に基づいてリモートセンターの上位装置もしくはディスクサブシステムが、更新データ反映処理を実行する。

【0006】また例えばデータベースをリモートコピー機能により二重化する場合では、二重化の対象となるデータの例としてはデータベースのデータ本体と、データ更新の履歴を記録するログデータがある。データ本体とログデータはそれぞれが密接に関連しており、データ更新のトランザクションが発生した際は、データベースはデータ本体の更新に加えログデータへ更新内容を書き込むことでデータ更新のトランザクションを完了し、データ内容の整合性を保証する。このような方法によりデータ内容の整合性を保証しているデータベースでは、信頼性の観点からデータ本体とログデータを異なるディスクサブシステムのボリュームに記録されるようにシステム設計が行われることが多い。そして、この様な、互いに異なるディスクサブシステムの各ボリュームに格納されたデータ本体とログデータをリモートコピー機能により二重化する場合においても、リモートセンターのボリュームにコピーしたデータ本体とログデータの内容は整合性を保った状態でなければならない。このためリモート

コピー機能ではメインセンターでデータ本体とログデータを異なるディスクサブシステムのボリュームに記録されている場合においても、メインセンターで更新された順序と同じ順序でリモートセンターにおいてもデータを更新する必要がある。これを実現するため、現状、リモートコピー実行の対象とするメインセンター内ディスクサブシステムが保有するボリューム(コピー元ボリューム)とリモートセンター内ディスクサブシステムが保有するボリューム(コピー先ボリューム)を一組のペアボリュームとし、ペアボリュームの集合を一つのグループ(以下、ボリュームグループと呼ぶ)として一括管理することで、リモートセンターでの更新順序を保持しデータ内容の整合性を保証している。

【0007】そして、メインセンター側からリモートセンター側へデータ本体やログデータを送信する際、公衆網を介して通信する場合がある。

【0008】

【発明が解決しようとする課題】従来技術によれば、メインセンターとリモートセンター間のデータ内容の整合性が必要なペア論理ボリュームの集合(ペア論理ボリューム群)は、一つのグループ(ボリュームグループ)として一括管理された制御を必要とする。このような制御方式において、例えばメインセンターとリモートセンターの各ディスクサブシステムを接続するためのデータ伝送路を形成している複数のデータ転送機器の一部を計画的な保守等により停止する場合、データ内容の整合性を保持するため、ボリュームグループ内の全てのペア論理ボリュームについてコピー元からコピー先へのコピーを一時停止した状態(以下、本状態をサスペンドと呼ぶ)にする必要がある。サスペンドにした後、停止していたデータ伝送路を再び起動することで、ボリュームグループ内ペア論理ボリューム群を再度二重化(以下、ペア再形成と呼ぶ)することができる。

【0009】しかしボリュームグループによる一括管理は、例えば複数のデータ伝送機器の内、一部の機器の停止であっても、内容一貫性の保持のため、二重化保持に影響のないペア論理ボリュームもサスペンドにし、一部の機器を再起動終了するまで、このサスペンドを継続していた。このためサスペンドが終了した後のペア再形成において、リモートコピーの構成として対象とする容量が大規模であるほど、ペア再形成時の二重化(コピー)が完了する時間に膨大な時間を要することになっていた。即ち、コピー元からコピー先へボリュームグループ内の各ペア論理ボリュームに属するデータ(大規模容量)を送付する際、サブシステム内の複数のデータ転送機器から並行してデータを送信しようとしても、これら複数のデータは、公衆網へ送出される段階で、複数のデータ転送機器の総容量より少ない容量の公衆回線に送出されるので、ペア再形成時の二重化が完了するまでに要する時間が膨大となっていた。

【0010】本発明の目的は、ボリュームグループ内のベア論理ボリューム群において、リモートコピー一時停止後、ベア論理ボリューム群のベア再形成時の二重化が完了する時間を短縮する方法を提供するものである。

【0011】

【課題を解決するための手段】上記目的は、第1記憶システムが有する複数の論理ボリュームのデータを、第2記憶システムへコピーする記憶システムにおいて、前記複数の論理ボリュームのうち、ある論理ボリュームのデータを、前記第2記憶システムへコピーを開始する第1ステップと、前記複数の論理ボリュームのうち、前記ある論理ボリューム以外の他の論理ボリュームのデータを、前記第1ステップに遅れて前記第2記憶システムへコピーを開始する第2ステップとを設けることによって達成される。

【0012】

【発明の実施の形態】本発明は、データ内容の整合性を保証したボリュームグループ内の全てのベア論理ボリューム群のコピーが一時停止した状態（サスペンド）になった場合、ボリュームグループ内の一部のベア論理ボリュームを選択し、ボリュームグループの中から順次ボリューム単位にベア再形成していくことを許可する二重化方式である。言い換えればサスペンドにした後、従来技術ではデータ内容の整合性からボリュームグループの全ベア論理ボリュームを一斉にベア再形成する必要があったが、本発明は、即ベア再形成が可能なベア論理ボリュームはすぐにベア再形成を実行し、またベア再形成が不可能なベア論理ボリュームについては不可能な理由が取り除かれた（例えばデータ伝送機器の復旧）後にベア再形成を実行する。これにより、ボリュームグループ単位で一括してベア再形成を実施する場合よりも、ベア再形成が不可能な理由が取り除かれた後のベア再形成を対象とする容量が少なくなるので、ボリュームグループ内ベア論理ボリュームの二重化が完了する時間を短縮することが可能である。

【0013】また本発明では順次ボリューム単位にベア再形成を実施することで、ベア再形成中にコピー先ボリューム間でデータ内容の一貫性はない。仮にベア再形成中にベア論理ボリュームのいずれかに障害が発生した場合、メインセンターのコピー元ボリュームのデータ復元にコピー先ボリュームを活用することはできない。これはデータベースを例にとると、データ本体を書き込んでいるコピー先ボリュームとログデータを書き込んでいるコピー先ボリュームとで時間をずらしてベア再形成を実行するため、両データのベア再形成を終了しないとデータ内容の一貫性はないことになる。これを回避するため本発明ではベア再形成を実施する前に、サスペンドにした時点のコピー先ボリュームの複製を、ディスクサブシステムがコピー先ボリュームとは別のボリュームに作成する。これにより、ベア再形成中にベア論理ボリューム

のいずれかに障害が発生した場合においても当該複製を利用することで、サスペンドにした時点までのデータを復元することができる。

【0014】以下、図面を参照しながら本実施形態について説明する。但し下記の説明に本発明が限定されるわけではない。

【0015】図1は、各々上位装置を装備する2つのデータセンターの間でデータの二重化を行うシステムの構成図である。

10 【0016】メインセンター101側のデータ記憶システムであるディスクサブシステム（MCU）102とリモートセンター103側のデータ記憶システムであるディスクサブシステム（RCU）104は、両センタ内の上位装置（ホスト）105、106を介さずに互いに接続され、MCU102が保有するデータをRCU104へ二重化を行うリモートコピーシステムを実現している。

20 【0017】メインセンター101においてMCU102は、インタフェースケーブル107を介してMCU102に対し、参照及び更新のデータ処理を行う中央処理装置（CPU）を持つ上位装置105に接続されている。また、MCU102は、上位装置105から参照及び更新処理を行うためのデータを格納する複数の一次ボリューム（P-VOL）108（108-1、108-2、・・・、108-n）を有する。

30 【0018】一方、リモートセンター103において、RCU104は、インタフェースケーブル110を介してCPUを持つ上位装置106に接続されている。この上位装置106は、メインセンター101の上位装置105が災害や故障等により本来の機能が果たせなくなった場合に、上位装置105の代替となって処理を行うことができる。また災害や故障時以外にも、RCU104に格納されているデータを利用して、メインセンター101の上位装置105とは異なる処理を上位装置105とは独立に実行することもできる。更に、RCU104には上位装置106から参照及び更新処理を行うためのデータを格納した複数の二次ボリューム（S-VOL）111（111-1、111-2、・・・、111-n）と三次ボリューム（T-VOL）112（112-1、112-2、・・・、112-n）を保有する。

40 【0019】メインセンター101内の上位装置105からMCU102が有するP-VOL（108-1、108-2、・・・、108-n）に対しデータの書込み指示があった場合、各P-VOLおよび各S-VOL（111-1、111-2、・・・、111-n）に対応する各インタフェースケーブル（109-1、109-2、・・・、109-n）を介して、リモートセンター103内のRCU104が有するS-VOL（111-1、111-2、・・・、111-n）へ向けて、送出される。この際、インタフェースケーブルから送出さ

れた書込みデータは、途中、回線多重分離かつ公衆網インタフェース(I/F)部130で多重され、公衆網上のある特定の公衆回線140経由でS-VOLへ向けて送出される。つまり、S-VOLへの書込みデータのコピーにおいては、全ての書込みデータが本公衆回線140上を通過することになる。従って、この公衆回線上に一度に大きなトラヒックがかかると、ここがボトルネックになることになる。なお公衆網I/F部130はインタフェースケーブルと1対1で接続されていても構わない。

【0020】本発明は、大きなトラヒックがこの公衆回線上に一度にかかるボリュームグループの一時停止後のベア再同期において、このトラヒックピークを時間的に分散し、ベア再同期にかかる時間を短縮するものである。後で、更に説明する。

【0021】図2は、MCU102の内部構成を示すものである。MCU102は、上位装置(ホスト)からのデータ授受を行うインタフェース制御部115と、上位装置から参照または更新されるデータを一時的に格納するメモリ116と、リモートコピーが一時停止中の更新データの格納位置に関する情報を格納するリモートコピー制御情報格納部117と、上位装置105のデータを記録する記録媒体としての磁気ディスクドライブ118、これらのデータのやり取りを制御するマイクロプロセッサ119、これらの要素を制御するディスクアレイサブシステム制御部120、それに加えてリモートコピーのコピーの実行状態の監視、および、リモートコピーをどのような設定にて行うかをユーザが設定できるサービスプロセッサパネル121を備える。ここで、磁気ディスクドライブ118は、MCU102には上位装置105から参照及び更新処理を行うためのデータを格納した図1に示す複数の一次ボリューム(P-VOL)108(108-1、108-2、・・・、108-n)を有する。

【0022】また図2のMCU102は、リモートセンタ103とデータの送受信を行うためのインタフェース制御部115を有する。そして、このインタフェース制御部115から、図1で示したインタフェースケーブル109-1、109-2、・・・、109-nが接続される。これらインタフェースケーブルは、回線多重分離・公衆網I/F部130に接続される。回線多重分離・公衆網I/F部130は、各インタフェースケーブルから受け取ったデータを、公衆回線140を介してリモートセンタ側の回線多重分離・公衆網I/F部131に送信する。リモートセンタ側の回線多重分離・公衆網I/F部131は、RCU104内のインタフェース制御部とインタフェースケーブルを介して接続される。この構成のため、メインセンタからリモートセンタへリモートコピーを行う場合、各サブシステム内では各論理ボリューム対応に互いに異なる通信路でデータ転送可

能であるが、公衆網上では全ての論理ボリュームに対する通信が1つの公衆回線上で行われることとなり、データ転送の負荷が集中する。本願発明では、特に負荷が集中するベア再同期時の負荷を時間的に分散し、ベア再同期に要する時間を短縮する方法を提供する。

【0023】なお、センタ間を結ぶインタフェースケーブル109には、例えばLED駆動装置によって駆動されている光ファイバリンクや、光ファイバケーブルを用いて一般にファイバチャネルと呼ばれるインタフェースプロトコルで駆動されるものである。また公衆回線140とはT3ネットワークやATMネットワークに代表される電気通信リンク、および公衆網I/F部130はチャネルエクステンダやファイバチャネルスイッチに代表されるインタフェースの接続距離を延長できるデータ転送機器である。従ってMCU102とRCU104の間には、途中に一般のファイバチャネルを接続したり、T3ネットワーク等と接続することも可能である。

【0024】図1に戻り、説明を続ける。MCU102は、上位装置105からMCU102が保有するP-VOL108への通常の参照及び更新処理と、P-VOL108のデータをRCU104が保有するS-VOL111へコピーする制御を行う。

【0025】具体的には、MCU102は、P-VOL108とコピー対象のS-VOL111の一組の論理ボリュームをベアとして、例えばP-VOL108-1とベアを成すのがS-VOL111-1、P-VOL108-2とベアを成すのがS-VOL111-2といったベア論理ボリュームの構成とベア論理ボリュームのコピー実行状態(ステータス)を管理する。RCU104は、MCU102から送信されたデータの書き込みの実行とベア論理ボリュームの構成およびステータスを管理する。

【0026】ここでステータスとは、P-VOL108とS-VOL111間のコピー実行状態を示すものであり、「デュプレックス」「サスペンド」の2つの状態である。「デュプレックス」は、P-VOL108とS-VOL111でベアの関係を維持した二重化の状態、言い換えればP-VOL108の更新処理を逐次S-VOL111へ反映する状態である。「サスペンド」は、ベアの関係を維持したままP-VOL108の更新データをS-VOL111へ反映する処理を一次中断した状態である。これらステータスは、上位装置上のアプリケーション113、114もしくはディスクサブシステム内のサービスプロセッサパネル121、ディスクサブシステムとLANで直接に接続されたコンソール上のアプリケーションからベア論理ボリュームに対して発行するコマンドによって遷移させることができる。

【0027】また本実施例ではデータ内容の一貫性を保持するボリュームグループをP-VOL108とS-VOL111のベア論理ボリュームとする。これによりベ

ア論理ボリュームがデュプレックスのステータスにある時、P-VOL108の更新順序とS-VOL111の更新反映の順序を一致させ、P-VOL108とS-VOL111のデータ内容に一貫性を保持させる。

【0028】さらに本実施例におけるRCU104が保有するT-VOL112は、ボリュームグループとして定義したP-VOL108とS-VOL111の全ペア論理ボリュームをサスペンドにした時点のS-VOL111の複製を格納した論理ボリューム群である。なお本実施例におけるS-VOL111の複製をT-VOL112に作成する技術は、ここでは詳細に記述しないが、リモートコピー機能がディスクサブシステム間でデータを二重化する技術であるのに対し、当該作成技術は同一ディスクサブシステム内でデータを二重化する公知の技術を利用したものである。本実施例では、S-VOL111と同一ディスクサブシステム内のT-VOL112を一組のペア論理ボリュームとしてS-VOL111のデータをT-VOL112に二重化したものである。

【0029】図4、図5（図4の続き）は、図1のリモートコピーの構成においてMCU102とRCU104によってデータを二重化している運用を行っている場合、インタフェースケーブル109と公衆網I/F部130で構成されている一部のデータ転送機器の保守等による予期せぬ障害、あるいは計画的な停止によって、二重化していたボリュームグループ内の全ペア論理ボリュームのコピーを一旦停止し、再びボリュームグループ内の全ペア論理ボリュームを二重化するまでの処理の手順を示す図である。この図の説明を通し、本発明である、コピーの一時停止後の全ペア論理ボリュームの二重化に必要な時間を短縮する手順を具体的に説明する。なお、本説明においては、各ペア論理ボリュームの状態を説明するために、必要に応じて図6および図7（図6の続き）を用いる。

【0030】まず処理を行う前提条件を説明する。データ転送機器（インタフェースケーブルを構成する）の停止前は、MCU102内のP-VOL108とRCU104内のS-VOL111がペア論理ボリューム、およびRCU104内のS-VOL111とT-VOL112がペア論理ボリュームとして形成され、各ペア論理ボリュームのステータスはデュプレックスであるとする。この状態では、上位装置105がMCU102の各P-VOL108に対して更新処理を行うと、当該更新データはリモートコピーによりRCU104内のS-VOL111へ反映され、かつRCU104内でS-VOL111を経由してS-VOL111とペアを成すT-VOL112へ反映されることになり、これら3つの論理ボリューム間のデータ内容は全て一致している。

【0031】また他の前提条件として、図1および図2で説明したように、MCU102とRCU104間のデータ転送に用いるインタフェースケーブル109は複数

存在し（109-1、109-2、・・・、109-n）、P-VOL108とS-VOL111のペア論理ボリュームにおいて、そのデータ転送はインタフェースケーブルで個々に独立しているものとする。例えば、P-VOL（108-1）とS-VOL（111-1）とのペアに対してはデータ転送機器であるインタフェースケーブル109-1が対応することになる。従って、各インタフェースケーブル109を構成するデータ転送機器の一部が保証等による計画的または障害により停止すると、P-VOL108とS-VOL111のペア論理ボリューム群の内、当該停止したデータ転送機器を含むインタフェースケーブルを使用していたペア論理ボリュームが、データ転送が不可能となり、このペア論理ボリュームに障害が発生することを意味する。なお、インタフェースケーブルを構成するデータ転送機器とは、インタフェースケーブル本体の他、公衆網I/F部130および図2には例示しないが、各インタフェースケーブル（109-1、109-2、・・・、109-n）に接続されるインタフェース制御部115に含まれるインタフェースケーブル対応の伝送制御回路部等が、これに該当する。

【0032】前提条件の説明が終わったので、次に、データ転送機器の停止後のペア再同期について、障害による予期しない場合と、計画的な停止による予期した場合に分けて、以下、具体的に述べる。

【0033】（1）障害による場合

各インタフェースケーブル109および公衆網I/F部130を構成している複数のデータ転送機器の一部に障害が発生（図4、ステップ201）した場合、MCU102ではRCU104からのデータ受領の応答を期待していたが応答がないことを検知し、RCU104ではMCU102からデータが送信されていないことを検知することによって、MCU102とRCU104では、ペア論理ボリュームの二重化が維持できないことを判断する。

【0034】次にMCU102はMCU102で管理するステータス管理テーブル（図6、図7）の中から、二重化の維持が不可能な当該P-VOL108の属性を「可」から「不可」に変更する（ステップ202）。ここでステータス管理テーブルとは、MCU102とRCU104が保持し管理するテーブルであり、図6、図7に示される。

【0035】図6、図7のMCU102のステータス管理テーブルでは、P-VOL108の論理ボリューム番号601（ディスクサブシステムの製造番号（図中の「0」）と当該ディスクサブシステムの論理ボリューム番号（図中の「0:01」、「0:02」））、当該P-VOL108の「デュプレックス」「サスペンド」の状態を示すコピー実行状態（ステータス）602、当該P-VOL108とペアを組むS-VOL111の論理

10

20

30

40

50

ボリューム番号（ペア相手先のディスクサブシステムの製造番号と当該ディスクサブシステムの論理ボリューム番号）603、当該P-VOL111のペア再形成の「可」「不可」を示す属性604を管理する。また図6、図7のRCU104のステータス管理テーブルは、S-VOL111の論理ボリューム番号605（ペア相手先のディスクサブシステムの製造番号（図中の「1」）と当該ディスクサブシステムの論理ボリューム番号（図中の「0:01」、「0:02」））、当該S-VOL111の「デュプレックス」「サスペンド」の状態を示すコピー実行状態（ステータス）606を管理する。

【0036】これらステータス管理テーブルの内容は、ユーザがペア形成を行う前に、予めMCU102にペア相手先のRCU104と当該RCU104配下の論理ボリュームを登録しておく。さらに属性とは、後述のボリュームグループ内の全ペア論理ボリュームがサスペンドになった後、ディスクサブシステムが自動的にボリュームグループ内のどのペア論理ボリュームをペア再形成するか決定するためのものであり、属性が「可」のものはボリュームグループ内の全ペア論理ボリュームがサスペンドになった後、すぐにペア再形成が可能なペア論理ボリューム、属性が「不可」のものはペア再形成ができないペア論理ボリュームであることを意味する。なお、図6、図7ではボリュームグループに2つのペア論理ボリュームが含まれるものとして説明する。

【0037】図4のステップ202で処理が完了した時点のペア論理ボリュームのステータス、およびMCU102とRCU104のステータス管理テーブルの状態は図6の区分（1）となる。

【0038】次にMCU102とRCU104は、データ転送機器の障害を検出すると、P-VOL108からRCU104への更新データの転送および反映を全て一時中断（サスペンドの状態）する（ステップ203）。なお、ボリュームグループとして定義したP-VOL108に関連するペア論理ボリュームの全てをサスペンドにするのは、障害が発生した一部のデータ転送機器を含む当該ペア論理ボリュームを停止し、当該ボリュームグループに含まれる他のペア論理ボリュームを動作させると、当該ボリュームグループに含まれる複数のペア論理ボリューム間の整合がずれ、ボリュームグループに含まれるペア論理ボリューム相互のデータ内容の一貫性を保持することができなくなるからである。つまり、一貫性を保持するため、ボリュームグループ内の全てのペア論理ボリューム群をサスペンドにし、P-VOLとS-VOLの整合性のずれが発生するのを押さえ、データの一貫性を保証している。更に、ここでは、障害発生時点でのS-VOLの内容がT-VOLに全て反映される。一方、ボリュームグループを定義していない場合、一貫性を保証する必要が無いため、データ転送機器が停止する

と、当該障害が発生したデータ転送機器を更新データの転送に使用していたペア論理ボリュームの状態をデュプレックスからサスペンドのステータスに移す。

【0039】また、ステップ203に対応して、それぞれ管理しているステータス管理テーブルのステータスをボリュームグループで定義していたP-VOL108およびS-VOL111に対して「デュプレックス」から「サスペンド」に変更する（図6の区分（2））。これにより、ボリュームグループ内の全ペア論理ボリュームはサスペンドになり、P-VOL108からS-VOL111への更新データの反映を中断し、P-VOL108とS-VOL111のデータ内容が一致した状態となる。結果、図4のステップ203で処理が完了した時点のペア論理ボリュームのステータス、およびMCU102とRCU104のステータス管理テーブルの状態は図6の区分（2）となる。

【0040】次にRCU104では、S-VOL111のステータスが「サスペンド」に変更されると、RCU104内のS-VOL111とT-VOL112によるペア論理ボリュームをサスペンドにさせる（ステップ204）。この時点のT-VOL112のデータ内容は、ボリュームグループの全ペア論理ボリュームをサスペンドにさせた時点のS-VOL111のデータ内容に一致している。またこれ以降では、S-VOL111のデータ内容に変更が発生しても、T-VOL112のデータ内容に変更されることはない。一方、図4のステップ204で処理が完了した時点のペア論理ボリュームのステータス、およびMCUとRCUのステータス管理テーブルの状態は図6の区分（3）となる。

【0041】次に、RCU104は、S-VOL111とT-VOL112によるペア論理ボリュームのステータスが全て「サスペンド」になった時点で、RCU104はMCU102にサスペンドが完了したことを報告し、MCU102はRCU104からの報告を受領する（ステップ205）。

【0042】報告を受領したMCU102は、ステータス管理テーブルの「可」の属性を持つP-VOL108を特定し、当該P-VOL108とS-VOL111とでペアを形成しているペア論理ボリュームに対してペア再形成を実施する（図5、ステップ206）。すなわちMCU102は、ステータス管理テーブル「可」の属性を持つペア論理ボリュームに対しステータスを「サスペンド」から「デュプレックス」に変更し、P-VOL108の更新データをS-VOL111に転送し、RCU104は更新データを受領し、S-VOL111に反映するとともにMCU102にデータを受領したことを通知する。従って、図5のステップ206の処理の最中、ペア論理ボリュームのステータス、およびMCUとRCUのステータス管理テーブルの状態は図7の区分（4）となる。即ち、この期間中、「サスペンド」と「デュプレックス」から「サスペンド」のステータスに移す。

レックス」の2つのステータスがボリュームグループ内のベア論理ボリュームに存在することになり、「デュプレックス」のステータスであるベア論理ボリュームのみP-VOL108の更新データをS-VOL111へ反映する。

【0043】また、この期間では、更新順序を保持するのは「デュプレックス」のステータスのベア論理ボリュームに関する更新順序が保持され、「サスペンド」のステータスのベア論理ボリュームについては、上位装置から更新データを受領するとS-VOL111へのデータ転送は行わず図2のリモートコピー制御情報格納部117で更新データの格納位置に関する情報を保持する。

【0044】その後、「サスペンド」のステータスであるベア論理ボリュームをベア再形成する時は、当該格納位置の情報を元に、サスペンドになってからP-VOL108が更新された分のデータをMCU102はS-VOL111へコピーを行う。この時MCU102はコピー中に上位装置から更新データを受領した場合は、前記「デュプレックス」のステータスの場合と同様、更新順序を保持した更新データの反映処理をコピーと並行して行う。これによりベア再形成を行うベア論理ボリュームの更新順序は保証される。

【0045】また、ステップ206の処理において、RCU104からのデータ受領の通知のないベア論理ボリュームがある場合は、ベア再形成を試みたが何らかの原因でRCU104へデータを転送できなかったことを意味するため、当該ベア論理ボリュームに関連する属性を「不可」に変更する(図5、ステップ207)。このステップ207に関するベア論理ボリュームのステータス、およびMCUとRCUのステータス管理テーブルの状態の変化は図示しない。

【0046】一方、ユーザは、ベア再形成中にアプリケーション113から定期的にボリュームグループ内のベア論理ボリュームのステータスと属性を監視し続け、属性「可」のベア論理ボリュームが全て「デュプレックス」のステータスであることを確認する(ステップ208)。これは、ユーザがステータスと属性の監視を続けるのは、データ転送機器の停止に影響のないベア論理ボリュームの全てが二重化されたかを確認するためである。

【0047】ユーザがこの状態を確認するためのアプリケーション113の画面表示を図8に示す。画面で表示される情報は、あるボリュームグループ(VG#=0001)における正当化時刻(00:00:00)と、MCU102のどこのP-VOL108とRCU104のどこのS-VOL111がベア論理ボリュームを構成しているかの構成情報、当該ベア論理ボリュームのステータスおよびベア再形成属性が表示される。図8の例では、ボリュームグループに4つのベア論理ボリュームが含まれており、ボリューム番号0:01および0:02

に関してはベア再形成属性が「可」かつステータスが「デュプレックス」であり、既に二重化が行われていることを示す。論理ボリューム0:03に関しては、ベア再形成属性が「可」でステータスが「サスペンド」であるから、これからステップ206が行われる状態を示している。つまり、この状態で、ディスクサブシステムは、ベア再同期が起動される。論理ボリューム0:04に関しては、ベア再形成属性が「不可」であるので、現状、このベア論理ボリュームが使用しているインタフェースケーブルが障害であることを示している。

【0048】なお、アプリケーション113は、定期的にベア再形成属性およびステータスを必要に応じてディスクサブシステムから収集し、図8に示すようなテーブルにして表示する機能を有するものとする。また、図8の正当化時刻とは、ホストから受け付けた書込みデータのうち、どの時刻までのデータがMCU側とRCU側の両方で一貫性がとれているかを示す時刻であり、この時刻までMCU側とRCU側のデータの一致性が取れていることを意味する。具体的には、MCUがホストから書込みデータを受けた時点で、書込み順序を示す順序番号(シーケンス番号)とその書込みを受け付けた時刻(タイムスタンプ)をつけて書込みデータ(コピーデータ)をRCUへ転送するが、RCUが、受信したコピーデータに添付された時刻のうち、順序番号順に跳びが無く連続に並べられたもののうち最新の順序番号を有するコピーデータに付与されている時刻が正当化時刻となる。従って、アプリケーション113は正当化時刻をみることで、どの時刻までのデータが、一貫性を持って二重化されているかを知ることができる。2つのMCUを用いる正当化時刻の考え方は後述する(図3利用)。

【0049】図5の説明に戻り、ユーザは、ボリュームグループ内のベア論理ボリュームがサスペンドになった原因を取り除く。すなわち障害となったデータ転送機器を所望の作業により障害を排除し、その後、再度起動する(図5、ステップ209)。

【0050】次にユーザは、アプリケーション113からボリュームグループ内のベア論理ボリュームの中から、前述まででベア再形成したベア以外で属性「不可」のベア論理ボリュームを全てベア再形成する指示を発行する。指示を受領したMCU102は、前述のベア再形成と同様にP-VOL108の更新データをS-VOL111に転送し、RCU104は更新データを受領し、S-VOL111に反映するとともにMCU102にデータを受領したことを通知する(ステップ210)。図5のステップ210で処理が完了した時点のベア論理ボリュームのステータス、およびMCU102とRCU104のステータス管理テーブルの状態が図7の区分(5)となる。

【0051】ユーザはアプリケーション113からボリュームグループ内の全てのベア論理ボリュームが「デュ

ブレックス」になったことを確認する。またMCU102は、ボリュームグループとして定義していたP-VOL108に関連した全てのペア論理ボリュームが二重化されたので、ステータス管理テーブルの属性を全て「可」に変更する。これによりボリュームグループの全てのペア再形成が完了したことになる(ステップ211)。

【0052】またユーザは、次のペア再形成の処理に備えて、アプリケーション114より、S-VOL11とT-VOL12とのペア論理ボリュームをペア再形成をMCU102に指示し(ステップ212)、MCU102からの指示に従いRCU104はS-VOL11とT-VOL12とのペア論理ボリュームをペア再形成を実行する。図5のステップ212で処理が完了した時点のペア論理ボリュームのステータス、およびMCU102とRCU104のステータス管理テーブルの状態は図7の区分(6)となるが、当該区分によるペア論理ボリュームのステータスは、データ転送機器を停止する前の図6の区分(1)の状態と同じになる。

【0053】以上のステップをまとめると、ステップ206の状態は、障害が発生したデータ転送機器を含むインタフェースケーブルを使用するペア論理ボリューム以外のペア論理ボリュームがペア再同期を再開した状態である。つまり、この状態は、インタフェースケーブルに障害が無いペア論理ボリュームに関するペア再同期は実行され、障害の有るインタフェースケーブルを使用していたペア論理ボリュームのペア再同期(ステップ210)は、まだ実行されない状態である。そして、従来、障害が回復した時点で、ボリュームグループ内の全てのペア論理ボリュームのペアを再同期する場合と比べて、インタフェースケーブルに障害が無いペア論理ボリュームに関するペア再同期と、障害の有るインタフェースケーブルを使用していたペア論理ボリュームのペア再同期とが、時間的なずれ(ステップ206とステップ210の時間的な差)をもって実行されることになる。従って、図1で説明した公衆回線140でのペア再同期に伴うデータ転送負荷が時間的に分散される。

【0054】このため、インタフェースケーブルに障害が無いペア論理ボリュームに関するペア再同期で送られるデータが、障害の有るインタフェースケーブルを使用するペア論理ボリュームのペア再同期で送られるデータより、時間的に早く送られているので、その分、従来の方式と比べて、ペア再同期のための時間が短くて済むことになる。言い換えれば、障害で無いインタフェースケーブルを使用するペア論理ボリュームのペア再同期は、障害が取り除かれた時点で、既に始まっているので、障害が取り除かれた後で全てのペア論理ボリュームの再同期を取るときと比べて、障害が取り除かれた後のペア再同期時のデータ転送量が少なくて済み、障害が取り除かれてからペア再同期が完了する、つまり、正常動作に

戻るまでの時間が短くなる。

【0055】また、ステップ204でS-VOLのサスペンドを終了後、S-VOLとT-VOLの間もサスペンドにし、これにより、障害発生時点でのボリュームグループのデータ状態を保持している。このため、例え、一部ペア論理ボリュームのペア再同期を行っている最中に新たな障害が発生して状態管理が正常で無くなり、P-VOLとS-VOLの関係が正常に保てなく修復不可能となっても、T-VOLのデータを使用して修復できるので、従来、障害が回復してから全ペア論理ボリュームの再開を行っていた場合と比べても、データ保証のレベルは保たれる。

【0056】以上が、インタフェースケーブル109および公衆網I/F部130を構成する複数のデータ転送機器の一部が障害によって停止した場合の処理である。次にデータ転送機器を計画的に停止する場合について説明する。

【0057】(2)定期的にデータ転送機器を停止する場合

まずユーザは、保守や構成変更により、データ転送機器を計画的に停止するため、アプリケーション113からボリュームグループ内の全ペア論理ボリュームをサスペンドにする指示を発行する(図4、ステップ214)。図4のステップ214で処理が完了した時点のペア論理ボリュームのステータスは図6の区分(1)となる。

【0058】サスペンドにする指示を受領したMCU102は、MCU102が管理するステータス管理テーブルのステータスを「デュプレックス」から「サスペンド」に変更し、RCU104に対してステータスを同様に変更するように指示する。RCU104はMCU102からの指示を受領すると、RCU104が管理するステータス管理テーブルのステータスを「デュプレックス」から「サスペンド」に変更する。これによりボリュームグループ内の全ペア論理ボリュームは、P-VOL108の更新データをS-VOL111への反映を中断したサスペンドの状態となり、P-VOL108とS-VOL111のデータ内容が一致した状態になる(図4、ステップ203)。ステップ203で処理が完了した時点のペア論理ボリュームのステータスは図6の区分(2)となる。そして、データ転送機器を計画的に停止するために、この時点でデータ転送機器を停止する(図4、ステップ215)。

【0059】次にRCU104では、S-VOL111のステータスが「サスペンド」に変更されると、RCU104内のS-VOL111とT-VOL112によるペア論理ボリュームをサスペンドにさせる。この時点のT-VOL112のデータ内容は、ボリュームグループの全ペア論理ボリュームをサスペンドにさせた時点のS-VOL111のデータ内容に一致している。またこれ以降では、S-VOL111のデータ内容に変更が発生

しても、T-VOL112のデータ内容が変更されることはない(図4、ステップ204)。図4のステップ204で処理が完了した時点のペア論理ボリュームのステータスは図6の区分(3)となる。

【0060】次にMCU102は、RCU104からS-VOL111とT-VOL112によるペア論理ボリュームがサスペンドになった通知を受領すると、ステータス管理テーブルの「可」の属性を持つP-VOL108を特定し、当該P-VOL108とS-VOL111とでペアを形成しているペア論理ボリュームに対してペア再形成を実施する。すなわちMCU102はP-VOL108の更新データをS-VOL111に転送し、RCU104は更新データを受領し、S-VOL111に反映するとともにMCU102にデータを受領したことを通知する(図5、ステップ206)。図5のステップ206で処理が完了した時点のペア論理ボリュームのステータスは図7の区分(4)となる。

【0061】ここでステップ206では、ボリュームグループ内の全てのペア論理ボリュームに対してペア再形成を実行するが、データ転送機器の停止によって二重化の維持が不可能なペア論理ボリュームについては、RCU104からの更新データ受領の通知はない。この場合MCU102では当該通知のないペア論理ボリューム、すなわちP-VOL108の更新データの転送を停止し、属性を「可」から「不可」に変更する(図5、ステップ207)。ただしRCU104から更新データ受領の通知があったペア論理ボリュームについては、更新データの転送を継続する。これによってボリュームグループ内の一部のペア論理ボリュームをペア再形成することが可能である。

【0062】次に、ユーザは、ボリュームグループから、計画的にサスペンドにしていたペア論理ボリュームをサスペンドから解除する動作を行う。すなわち、計画的に停止していたデータ転送機器を再度起動する(図5、ステップ209)。

【0063】次にユーザは、データ転送機器が再度起動された事を確認すると、アプリケーション113からボリュームグループ内のペア論理ボリュームの中から、前述まででペア再形成したペア以外で属性「不可」のペア論理ボリュームを全てペア再形成する指示を発行する。指示を受領したMCU102は、前述のペア再形成と同様にP-VOL108の更新データをS-VOL111に転送し、RCU104は更新データを受領し、S-VOL111に反映するとともにMCU102にデータを受領したことを通知する(図5、ステップ210)。図5のステップ210で処理が完了した時点のペア論理ボリュームのステータス、およびMCUとRCUのステータス管理テーブルの状態は図7の区分(5)となる。

【0064】ユーザはアプリケーション113からボリュームグループ内の全てのペア論理ボリュームが「デュ

プレックス」になったことを確認する。またMCU102とRCU104は、ボリュームグループとして定義していたP-VOL108とS-VOL111による全てのペア論理ボリュームが二重化されたので、ステータス管理テーブルの属性を全て「可」に変更する。これによりボリュームグループの全てのペア再形成が完了したことになる(図5、ステップ211)。

【0065】またユーザは次のペア再形成の処理に備えて、アプリケーション114より、S-VOL111とT-VOL112とのペア論理ボリュームをペア再形成する(図5、ステップ212)。図5のステップ212で処理が完了した時点のペア論理ボリュームのステータス、およびMCU102とRCU104のステータス管理テーブルの状態は図7の区分(6)となるが、当該区分によるペア論理ボリュームのステータスは、データ転送機器を停止する前の区分図6の(1)の状態と同じになる。

【0066】以上説明したように、計画的にデータ転送機器を停止した場合でも、障害によるデータ転送機器の停止の場合と同様に、計画的に停止したデータ転送機器を含まないインタフェースケーブルを使用するペア論理ボリュームに関しては、計画的に停止したデータ転送機器を再動作させる前から、ペア再同期を始めるので、計画的なデータ転送機器の停止が終了した後から全てのペア論理ボリュームの再同期を取るときと比べて、計画的停止が終了した後のペア再同期時のデータ転送量が少なく済み、計画的データ転送機器の停止が終了してからペア再同期が完了するまでの時間、つまり、正常動作に戻るまでの時間が短くなる。

【0067】以上が、インタフェースケーブル109および公衆網I/F部130を構成している一部のデータ転送機器の保守等による予期せぬ障害、あるいは計画的な停止によって、二重化していたボリュームグループ内の全ペア論理ボリュームのコピーを一旦停止し、再びボリュームグループ内の全ペア論理ボリュームを二重化する処理の説明である。

【0068】なお本実施例では詳細に説明しないが、ユーザのアプリケーションによる指示は上位装置だけでなく、ディスクサブシステム専用の管理コンソールから指示する実施形態でも構わない。またデータ転送機器を計画的に停止する場合は、ユーザが当該機器の停止によりペア再形成の不可能なペア論理ボリュームを特定し、当該ペア論理ボリュームに対し属性を「不可」に変更する指示をアプリケーションから与えても構わない。さらに本実施例によるリモートコピー構成は、図1に示したようにメインセンター101側のMCU102とリモートセンター103側のRCU104が1対1に対応して接続されるシステム構成についての実施形態の構成であったが、他の実施形態として、複数台のMCU102が1台のRCU104に対応している場合、もしくは逆に1

台のMCU102が複数台のRCU104に対応している場合、または複数台のMCU102と複数台のRCU104がそれぞれ対応している場合、いずれの構成においても本実施例によるリモートコピー処理が適用できる。

【0069】複数台のMCU102が、1台のRCU104に対応している場合は、ボリュームグループのデータ更新の順序性を保持するための仕組みが複雑になるので、図3を用いて説明しておく。MCU102は、上位装置105から更新データを受領すると、当該更新データにタイムスタンプとシーケンス番号を付与して上位装置105の更新指示によるデータ書き込みとは非同期にRCU104に送付する。RCU104では受領した更新データの順序が必ずしもシーケンス番号の順序とは一致しないため、シーケンス番号の順序に並び替えた後、RCU104内のメモリに更新する。またRCU104では受領した更新データを送り元のMCUごとに管理する。この管理の基準は例えばディスクサブシステムの製造番号であり、各MCUを特定できるものである。

【0070】次にRCU104では、各MCUごとにシーケンスが保証されているデータの中から最新のデータの時刻値（タイムスタンプ）を求める。図3の例ではMCU#1の最新のデータの時刻値はT7、MCU#2の最新のデータの時刻値はT5となる。そして各MCU間で最新データの時刻値を比較し、最古の時刻を正当化時刻（更新順序の整合性が保持されている時刻）と決定し、当該時刻以前のデータを磁気ディスクドライブのS-VOL111に反映する。図3の例では正当化時刻がT5となり、T5以前のデータをS-VOL111に反映することになる。またRCUが複数の場合も、各RCU間で最新データの時刻値を比較し最古の時刻値を正当化時刻とする。

【0071】以上本発明を纏めると、ボリュームグループとして定義されたベア論理ボリューム群の中で、データ伝送機器の停止後も二重化保持に影響のないベア論理ボリュームは当該機器の停止中も二重化を保持し続け、当該機器停止によって二重化保持が不可能なベア論理ボリュームだけを当該機器の復旧後に二重化（ベア再形成）させる。これにより当該機器復旧後に二重化させる容量は、ボリュームグループ内の全ベア論理ボリュームの容量よりも少なくなるので、ボリュームグループ内の全ベア論理ボリュームのベア再形成を完了するまでの時間は、ボリュームグループ内の全ベア論理ボリュームを一括してベア再形成する場合よりも短縮され、早期二重化が可能となる。更に、図3で示した更新手順を用いることで、送信側MCUが複数になった場合でもそのデータの一貫性を保証した形で、データ伝送機器が止まった後のベア再形成時間を短縮することができる。この際、正当化時刻を用いることで、ボリュームグループが保証できるデータ時刻も知ることができる。

【0072】また、上記説明ではデータをリモートサブシステムに送る公衆回線がデータ転送遅延の障害になるとして説明しているが、複数のインタフェースケーブルから送出されるベア再形成用のデータがメインセンターからリモートセンターへ転送される間で、何らかの形で、そのデータ転送能力を制限される状態（例えば、他のリモートコピー・トラヒックによりデータ転送能力が制限される、転送品質が良くなり再送等が頻発し転送自体に多くの時間を有する、等）を有する伝送路であれば、本願発明を適用することによって、即ち、ベア再形成するタイミングをずらすことによって、ベア再形成する時間を短縮することができる。

【0073】また、上記説明では、データ転送機器が停止していないベア再形成の開始をユーザが指示していたが、MCUおよびRCU内のマイクロプロセッサが、ステータス管理テーブルの内容を判断して、自動的にベア再形成を指示するようにしても良い。

【0074】

【発明の効果】リモートコピー機能において、ボリュームグループ内のベア論理ボリューム群を順次ボリューム単位にベア再形成を実施することで、ボリュームグループ内の全ベア論理ボリュームのベア再形成を完了するまでの時間を削減することができる。

【図面の簡単な説明】

【図1】実施例におけるリモートコピーシステムの構成概要を示す。

【図2】実施例におけるディスクサブシステムの構成を示す。

【図3】正当化時刻決定の概念を示す。

【図4】データ転送機器の停止時における処理フローを示す。

【図5】図4の続きである。

【図6】各ステップ処理のベア論理ボリュームのステータスと、当該ステータス管理テーブルの遷移状態を示す。

【図7】図6の続きである。

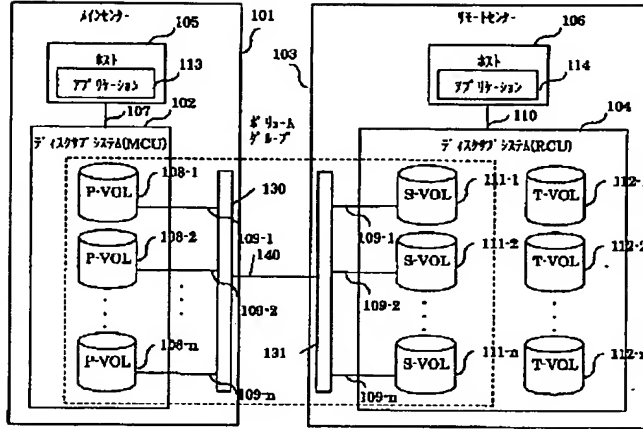
【図8】アプリケーションの画面表示例を示す。

【符号の説明】

101：メインセンター、102、104：ディスクサブシステム、103：リモートセンター、105、106：上位装置、107、109、110：インタフェースケーブル、108：一次ボリューム、111：二次ボリューム、112：三次ボリューム、113、114：アプリケーション、115：インタフェース制御部、116：メモリ、117：リモートコピー制御情報格納部、118：磁気ディスクドライブ、119：マイクロプロセッサ、120：ディスクアレイサブシステム制御部、121：サービスプロセッサパネル、109-1、109-2：インタフェースケーブル。

【図1】

図 1



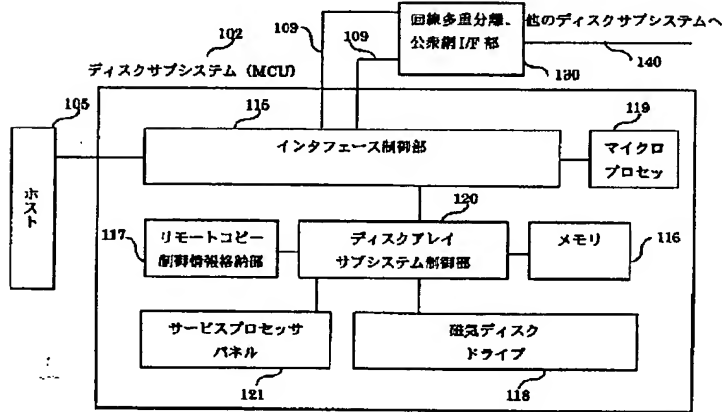
【図8】

図 8

VG# = 0001		正当化時刻 = 00:00:00	
P-VOL(MCU#-VOL#)	S-VOL(RCU#-VOL#)	ステータス	ペリ外形成属性
0-0-01	1-0-01	デュプレックス	可
0-0-02	1-0-02	デュプレックス	可
0-0-03	1-0-03	デュプレックス	可
0-0-04	1-0-04	デュプレックス	不可

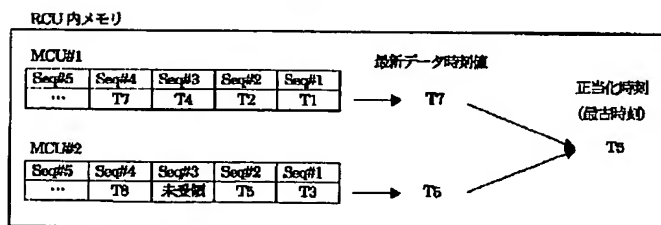
【図2】

図 2



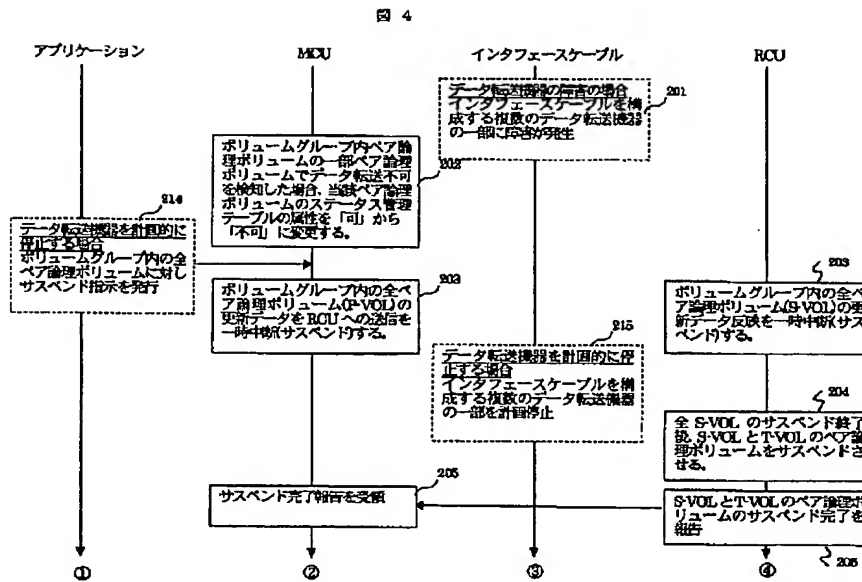
【図3】

図 3

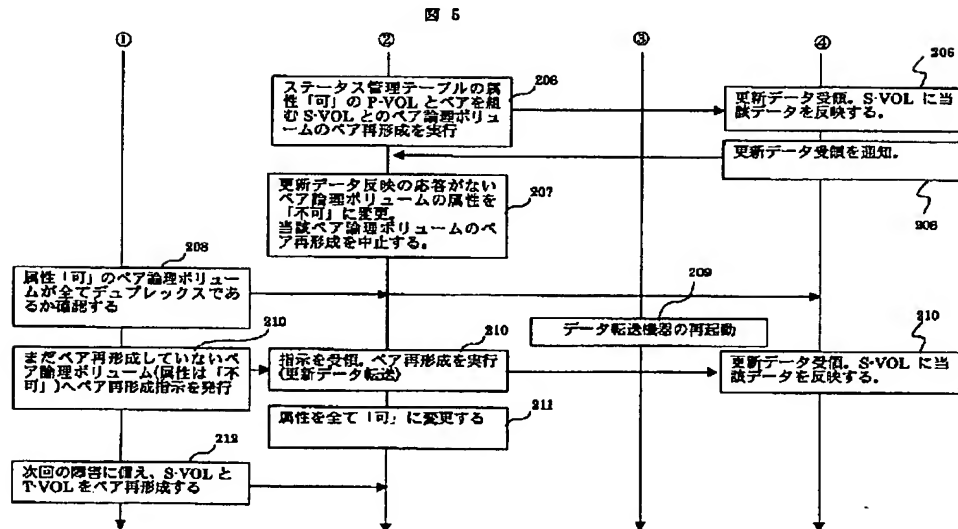


Seq# : シーケンスの番号
Tx : タイムスタンプ

【図4】



【図5】



【図6】

図 6

区分	ペア論理ボリュームの ステータス遷移			MCUステータス管理テーブル				RCUステータス 管理テーブル	
	P-VOL	S-VOL	T-VOL	P-VOL	P-VOL ステータス	S-VOL	ペア再形 成属性	S-VOL	S-VOL ステータス
(1)				0-0:01	デュープレックス	1-0:01	可→不可	1-0:01	デュープレックス
				0-0:02	デュープレックス	1-0:02	可	1-0:02	デュープレックス
(2)				0-0:01	クォータ	1-0:01	不可	1-0:01	クォータ
				0-0:02	クォータ	1-0:02	可	1-0:02	クォータ
(3)				0-0:01	クォータ	1-0:01	不可	1-0:01	クォータ
				0-0:02	クォータ	1-0:02	可	1-0:02	クォータ

601 602 603 604 605 606

【図7】

図 7

区分	ペア論理ボリュームの ステータス遷移			MCUステータス管理テーブル				RCUステータス 管理テーブル	
	P-VOL	S-VOL	T-VOL	P-VOL	P-VOL ステータス	S-VOL	ペア再形 成属性	S-VOL	S-VOL ステータス
(4)				0-0:01	クォータ	1-0:01	不可	1-0:01	クォータ
				0-0:02	デュープレックス	1-0:02	可	1-0:02	デュープレックス
(5)				0-0:01	デュープレックス	1-0:01	不可→可	1-0:01	デュープレックス
				0-0:02	デュープレックス	1-0:02	可	1-0:02	デュープレックス
(6) (1)に 同じ				0-0:01	デュープレックス	1-0:01	可	1-0:01	デュープレックス
				0-0:02	デュープレックス	1-0:02	可	1-0:02	デュープレックス

601 602 603 604 605 606

フロントページの続き

(72)発明者 島田 朗伸
神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内

(72)発明者 中野 俊夫
神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内

Fターム(参考) 5B018 GA04 HA04 MA11
5B065 BA01 CA11 CC08 CE22 CS01
EA31 EA35 EA38 ZA01
5B082 DC05 DE04